NHII and Clinical Research White Paper 2004 Annual Meeting

Harold P. Lehmann, MD PhD Peter L. Elkin, MD

Steven H. Brown, MD MS Kimberly Elenberg, MS Reed M. Gardner, PhD Charles Jaffe, MD PhD Michael G. Kahn, MD PhD Stephen Rosenfeld, MD

Corresponding author: lehmann@jhmi.edu

Executive Summary

Clinical research is an engine that advances medical care and creates new business opportunities. As the nation's health care becomes increasingly computerized, the opportunity exists to detect, develop, and disseminate new medical insights throughout the healthcare environment rapidly but only if specific needs and issues are acknowledged and incorporated into the national health information infrastructure. This document lays out the needs from multiple perspectives and at multiple layers, taking into account current needs and future directions.

Table of Contents

Executive Summary	1
Table of Contents	
From NHII 2003	2
2004	
Consensus Recommendations, 2003	2
Short term	
Medium term	
Future status	2
Progress made in the past year	3
Consensus Recommendations, 2004 < To be filled in at the meeting!>	
Short term	
Mid term	
Long term	3
Framework for breakout discussion	
Introduction	3
I. Stakeholders' Perspectives	
II. Research Function	
III. Research Workflow	
IV A. Research Information System Needs	7
IV B. Healthcare Delivery System Needs	
V A. Data, Information, and Knowledge/Repository Needs	
V B. Algorithmic/Tool Needs	
VI. Technology	10

Goal State

From NHII 2003

The 2003 Breakout session combined clinical- and population-level research: (Posted by Shawn Murphy, MD PhD)

- Multiple purpose data use
- Research and population health data standardization
- Reliable means of linking people in various data sets
- Support for secure access and information exchange
- Support for various ownership models of data

2004

This year's charge was to focus more on clinical research

- Support for the entire research life cycle
- Support for research throughout the patient's life
- Support for research in all clinical venues, including remote and home-based investigations
- Support for research at multiple levels of abstraction, from molecular-biological through population
- Support for human subjects safety, dignity and decision-making of research
- Promote single system for clinical care and clinical research while acknowledging the diversity of clinical research settings and populations
- Support sharable modules of research workflow, decision-support, monitoring and quality control
- Patient-specific application of research results

Consensus Recommendations, 2003

Short term

Adopt a complex adaptive system approach to support data exchange and interoperability in research and population health. Define the key elements to allow for this kind of approach: HL7

Establish standard set of patient characteristic linker variables and accompanying logical methods for matching patients for research and population health studies. Match can tolerate <100% accuracy: AHRQ, NIST, CDC

Establish national registry of data definitions, data sets, and metadata for research and population health: HL7, LOINC, NLM (UMLS), NIH

Develop incentives for data providers to conform to NHII standards and make data available for research and population health (e.g. Make participation required to receive funding, distribute direct financial incentives, allow equal access to data if they contribute): NIH

Develop incentives for data providers to conform to NHII standards and make data available for research and population health (e.g. Make participation required to receive funding, distribute direct financial incentives, allow equal access to data if they contribute): NIH

Medium term

Develop procedures, processes, and guidelines to ensure that the research community will communicate knowledge back to the public to improve personal health decisions: AHRQ, CDC, NIH

Funding and leadership for research and population health in NHII should come from state and local governments and private sector, as well as federal agencies.

Future status

Increase data exchange and interoperability.

Shorten research cycle.

Allow focus to shift away from data collection and onto data analysis.

Return population based insights to consumers

Progress made in the past year

- Identification of networks as an area for development in the NIH roadmap
- Endorsement of several critical CHI standards (only immediately applicable to healthcare in the federal system, but meant to lead by mass action/demonstration,
- Increased visibility of clinical research in this whole process (perhaps thanks to the roadmap?)
- Public Health Data Standards Consortium (<a href="http://phasea.org/phasea.

Consensus Recommendations, 2004 < To be filled in at the meeting!>

The recommendations are based on two principles: NHII should support the tasks researchers (and other stakeholders) currently have to perform; NHII should support the tasks researchers (and other stakeholders) will have to perform in the near future.

Short term

Mid term

Long term

Framework for breakout discussion

- The NHII will benefit from research
- Research will benefit from NHII
- The NHII must acknowledge the real needs of multiple perspectives with regards to research
- The NHII should be designed to support higher-level functions, based on research workflow and practice
- NHII must support Translational Research
- The NHII should anticipate functions that are not yet demanded
- The NHII process should avoid putting in place barriers that engender conflict between clinical care and research
- The NHII process must protect intellectual property

Introduction

Research fuels evidence based medical practice and advances patient care. Research innovation keeps American medical care to a high standard of excellence. "Research" covers a wide range of levels of evidence, from population-based observational studies through formal clinical trials to evidence syntheses, and evaluates a wide range of technologies and interventions, from innovative pharmaceuticals through devices through new practice models through population-based interventions, to information technology interventions for a wide range of purposes, such as surveillance, risk assessment, and intervention assessment.

Research will benefit from an NHII that makes it easier to assemble quickly assemble both small, highly-defined populations and large, broad-based populations to study, that makes it easier to share data across studies, and that makes it easier to answer important questions that do not currently get addressed due to infrastructure issues, such as optimal treatment strategies, long-term chronic disease care, or low-probability risks.

NHII will benefit by being required—in supporting clinical research—to develop and to deploy technologies that "work" in extremely diverse clinical practice settings, including patient homes, require data sharing across diverse institutions, and across long periods of time (chronic disease research), while permitting merging and combining of disparate data into a logical and analyzable whole. NHII will further benefit from research that pushes the limits of integrating bioinformatics with clinical data and in applying evidence-based interventions to larger regions.

¹ Report of the Ad Hoc EHR-Public Health task is at http://phdatastandards.info/about/committees/ehrph.htm)

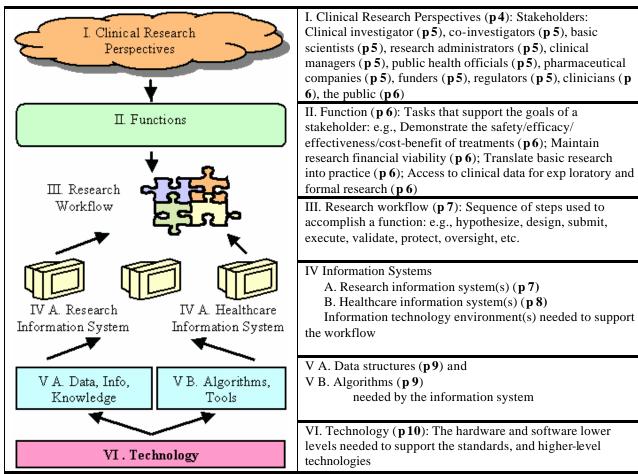
Innovation requires information to assess the unmet medical needs in current practice. Innovation requires information to assess the potential value of tentative/initial/small-scale approaches that are based on more basic research hypotheses. Innovation requires information to adequately understand the risks and benefits of a promising new medical advance on specific, tightly controlled populations. Innovation requires information to adequately understand societal implications of a broadly-deployed new intervention. And innovation needs information to know how the new innovation has not completely solved the original unmet need.

As the NHII grows, it will play an important role with regard to research. That role can be enabling or inhibiting, depending on the extent to which the growth process pays attention to the interests of stakeholders in research. As an example, the regulatory burden around clinical trials is very high, and falls primarily on the investigators. There are concrete examples including discordant coding standards (MedDRA for encoding adverse events and ICD-9CM for encoding reimbursement diagnoses, neither are adequate for encoding most clinical events), the lack of incentives to appropriately fund and support IRB infrastructures, etc. (the system is been simultaneously designed to contain costs while placing increasing unfunded mandates on research). These conflicts are explicitly recognized by the NIH roadmap, which has a track on regulatory simplification and rationalization. The goal of this White Paper is to provide a common consensus on issues involved.

The goal of these breakout sessions are to explore the needs of the clinical research community, broadly defined, and to specify functionality needed of the NHII to support those needs.

This document takes an informatics approach by dividing the issues into informatics-based layers of abstraction (see **Table 1**). Within each layer, multiple stakeholder perspectives will be considered; available and needed standards will be listed, as will be metrics for assessing the size of current gaps and for monitoring progress as the NHII is realized.

Table 1. Layered framework for considering clinical research needs and the NHII.*



^{*}Standards (the arrows or thick lines between layers) and metrics are needed at each level; they are not included explicitly in this table.

I. Stakeholders' Perspectives

Clinical investigators have exacting information needs, whether they are single investigators or part of a multicenter team. Just in terms of study execution (after study design), they need to negotiate the process of study approval by Institutional Review Boards (IRBs) for human-subject research. They then need to recruit patients into their studies and to collect data that is more detailed, more complete and more validated than general clinical data. In addition, they need to track outcomes in their patients, especially adverse outcomes, with a degree of reliability and completeness that is, again, greater than in general clinical practice. They need to report on their process to funding agencies, again, in greater detail than is found in clinical practice.

Clinician participants (co-investigators) increasingly are not working at academic settings, yet they have the same information needs. They need mechanisms to weave together the research process and the care process for optimal efficiency. There are vendor solutions, home grown technologies (developed by academia and pharma), and government systems (NIH/NCI). How do these relate to the NHII framework? What can we contribute? How will the road to success look different from the perspective of academia and pharma and government? There are broad differences in purpose and process between these two perspectives.

Basic scientists increasingly look to the clinical population as a source of hypotheses, as a source of research material, and as a target to test out translational ideas. While much has been written³ on the potential for bioinformatics information to affect or even to swamp the electronic patient record environment, it will be in academic research settings where these issues will first be encountered.

The terms "basic scientist" and "clinical investigator" may no longer represent a clear distinction in a world where multi-disciplinary and inter-disciplinary researchers are necessary to accomplish the needed work.⁴

Research administrators are responsible for ensuring that research performed at the institution follows regulations governing the research, as well as ensuring appropriate use of funding.

Clinical managers are not traditionally thought of as clinical researchers, but, because of their need to extrapolate from historically collected data on a local population of patients, they function as researchers. They increasingly use the patient record for evaluating the quality of care, 5 and, conversely, the availability of electronic records has made this type of "research" more common. There is a need for tools that help clinical managers, much as executive decision support tools help middle-level industrial managers.

Public health officials, while focused, generally, on population-level data, often use analytic techniques developed for the analysis of clinical research data.⁶

Pharmaceutical companies conduct or fund a large portion of American clinical research. While project-oriented needs of pharmaceutical researchers are similar to those of clinical investigators, the former also need the ability to link studies together and to provide a view of entire research program, as opposed to the single study. This is especially true if large-scale research programs are to take advantage of novel methods, such as Bayesian design. Pharma has initiatives regarding FDA submission, orchestrated through HL7 and CDISC. There is almost no effort to become integrated within healthcare IT.

Funders' primary need is to establish that important and methodologically rigorous studies are supported.

Regulators are concerned about the safety and efficacy of medical treatments. Regulators are charged with ensuring that all research is conducted within an approved legal framework and that the data provided as "evidence" for a new therapeutic drug, device, or system of care is acceptable for widespread dissemination. Regulators require access to detailed trial information to make independent assessments for approving a trial and/or for accept trial conclusions.

² Department Of Health And Human Services, National Institutes Of Health, Office For Protection From Research Risks. Part 46, Protection Of Human Subjects. Revised November 13, 2001

Effective December 13, 2001. http://ohrp.osophs.dhhs.gov/humansubjects/guidance/45cfr46.htm)

³ Friedman CP, et al. Training the next generation of informaticians: the impact of "BISTI" and bioinformatics--a report from the American College of Medical Informatics. J Am Med Inform Assoc. 2004 May-Jun;11(3):167-72. Epub 2004 Feb 05.

⁴ Tabak L, Grady P. Research teams of the future. In: NIH Roadmap for medical research: A briefing by the Director and senior staff; 2004 Feb 27, 2004; Bethesda, MD; 2004. http://nihroadmap.nih.gov/briefing/

⁵ Henry SB. Informatics: essential infrastructure for quality assessment and improvement in nursing. J Am Med Inform Assoc. 1995 May-Jun;2(3):169-82.

⁶ Adams T, et al. Lessons from the central Hampshire electronic health record pilot project: issues of data protection and consent.BMJ. 2004 Apr 10;328(7444):871-4. Inoue LY, Thall PF, Berry DA. Seamlessly expanding a randomized phase II trial to phase III.Biometrics. 2002 Dec;58(4):823-31.

Clinicians who read the results of research need to understand the level of evidence provided them by particular research, as well as the level of importance and the relationship to other evidence.

The public must be able to verify that their medical care is evidence-based. This is especially so for disease-specific advocacy groups and issue-specific non-profit agencies

STANDARDS required at this level: METRICS required at this level:

II. Research Function

Overall Goal: to move in the direction from:

Data \Rightarrow Information \Rightarrow Knowledge \Rightarrow INTELLIGENCE \Rightarrow ACTION

In 2004, we are still collecting data as a "one of" experiment. When possible we analyze the data and form conclusions from our research trials. In order for this information to be stable, multiple trials need to be performed verifying these results. In the future we envision seeing multi-center designs with the rigor to settle questions definitively. The Information generated by trials must be put into practice to be considered Knowledge. This takes a translational effort which to date takes on average seventeen years to occur. Most would agree that this timeline is too long to support the rapid growth that we know will be needed in the age of the genomics revolution. Turning this Knowledge into Intelligence is harder and requires significant Informatics support. One must learn to find synergy between knowledge and knowledge sources. This will enable the discovery of best practice. Best practice includes both efficacy and safety be taken into consideration. Interactions between physiology, multiple diseases and multiple therapies linked to the patient's genetic fingerprint can and will pose a significant combinatorial problem that requires a merger of complexity theory and iterative design principles to solve. Future clinical research intelligence will ensure that our children and grandchildren will have safer and more efficacious contact with the healthcare system.

Demonstrate the Efficacy and Effectiveness of Clinical Treatments

The core purpose of clinical research is to establish the efficacy and safety of new treatments, as well as to demonstrate the effectiveness of new treatment strategies against current standards.

Business of Clinical Research: Maintain financial viability of clinical research

Workflow, workflow, workflow. This should be the mantra for clinical research systems. The workflow for the development of a methodologically rigorous and appropriate trial. The workflow for approving the trial by the IRB and funding agencies. The workflow for executing the trial. This includes support for comparable data, sharing data across sites and aggregating, analyzing and reporting the data and its associated knowledge. Some areas of specific interest are sample handling, lab book solutions, database design and maintenance solutions, and animal handling solutions. We envision this reporting to be in traditional print / digital media and also as knowledge input to clinical research data warehouses. Implicit in this vision is a set of intelligent agents designed to mine the data and fuel clinical practice and further hypothesis driven research.

Support for Translational Research: Translate basic research into clinical practice

Seventeen years is too long of a delay to translate research data into the practice of medicine. We must speed the time from discovery to implementation of this knowledge into the practice of medicine. Research into the preferred methods for bridging this gap should be supported by the NIH. Only through rigorous study and experimentation will we find the solutions needed to decrease this egregious delay.

Repurposing (e.g. Data Mining) Clinical Data for Research

The ability to begin to build a storehouse of data, which once collected, can serve as the substrate for future study is a primary goal of knowledge representation. Clinical research is expensive and time consuming. An information infrastructure, which uniformly uses standard mechanisms to represent clinical research data, will serve to decrease costs and some of the delays in recruitment associated with clinical research. With comparable data as our infrastructure, we can more safely combine data from multiple sites, increasing our sample sizes from the thousands to millions of patients in many of our trials that deal with common illness. For rare illnesses, we will be able to provide evidence where previously recruitment for a randomized controlled trial was not feasible. Once accumulated, intelligent agents can be crafted which would constantly be interrogating the world's clinical research data looking for associations and trends that have previously not been reported. Post-marketing surveillance would

⁸ Contopoulos-loannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications. Am J Med. 2003 Apr 15;114(6):477-84.

have the capability to look at each and every person taking a new drug, and in that way rapidly identify adverse drug events.

Summary Statement: By the year 2010, all US research data will be stored in a comparable format. That data assumptions and specific aims from experiments will be made explicit and computable. Thus enabling us to maximize the reusability of data and facilitate large clinical trials including post-marketing surveillance. Further, this computable research data will be exposed in such a way, that Informaticians interested in education can rapidly deliver real time teaching to clinicians at the point of care, improving patient safety and moving us closer to Informatics empowered personalized medicine.

Clinical investigator Basic scientist Research administrator Clinical manager

Public health officials have many concerns. Standard concerns involve the need public health surveillance in a cost-effective and comprehensive manner. Recent concerns regarding biodefense preparedness only highlights this standard concern. An emerging need is genetic epidemiology—How can we collect "clinical data" to allow us to use genealogical/genetic data to better understand how other factors such as medical care and the environment can help us with determining the causes of disease.

Pharmaceutical companies Funders Regulators Clinician-reader Public-reader

STANDARDS required at this level: METRICS required at this level:

III. Research Workflow

Clinical investigator: Clinical trial planning and design, patient screening, eligibility and recruitment, IRB status management, study monitoring, patient scheduling, budgeting and milestones, adverse event reporting, case report form and study execution, workflow configuration, and integration with internal and third party systems.⁹

Basic scientist: Communicate non-CLIA-approved results to patients or to the clinical team; communicate genome/proteome results to patient/clinical team.

Research administrator: Requires a paperless, electronic method to enable investigators to submit, and for investigators and IRB committee to track, and to review the scientific, regulatory, and compliance information required for the safe conduct of human subjects research. The system provides a platform for the IRBs and other research compliance committees together with the local research community to share critical information regarding the submission and review of new applications, amendments, protocol events, and continuing reviews. ¹⁰

Clinical manager
Public health official
Pharmaceutical companies
Funder
Regulators
Clinician-reader
Public-reader

STANDARDS required at this level: METRICS required at this level:

IV A. Research Information System Needs

Information Infrastructure for Clinical Research

⁹ Velos, Inc. http://www.velos.com/velos/topten.html (Jan 27, 2004)

¹⁰ Johns Hopkins eIRB system. http://www.hopkinsmedicine.org/webnotes/human/0406.cfm (June 2004)

To ensure comparable data within, among and between experiments and institutions we need a standard, robust data infrastructure capable of supporting and representing the data needed for clinical research. This includes compositional controlled health terminologies, which are capable of representing the diverse types of clinical research information. Ontologies, which represent the business of clinical research, need to be created and maintained to assist Informaticians in their efforts to create systems, which are capable of representing this knowledge. This is an imperative for systems that need to implement support for the workflow needs of clinical research. We need mechanisms for converting legacy data to this standard format. Therefore interfaces are needed between these research databases and master patient indices, scheduling systems, EHRs, Laboratory Systems, Pathology Systems, Radiologic Systems (indeed Media Asset Management systems in general), billing systems, Nursing records, Order Entry Systems and Pharmacy Systems. These interfaces need standard HL7 messages to ensure comparability of data across systems and environments. A strong information infrastructure is the support we can count on as we build the clinical research systems of the 21st Century.

To what degree does the proposed National Electronic Clinical Trials/Research Network (NECTAR)¹¹ support these needs?

Clinical investigator Basic scientist Research administrator Clinical manager Public health official Pharmaceutical companies must support trials conducted in multiple setting: academic and non-academic, yet with the same technology requirements. There are vendor solutions, home grown technologies (developed by academia and pharma), and government systems (NIH/NCI). **Funders** Regulators Clinician-reader

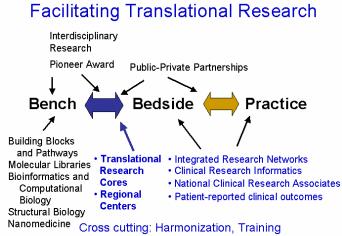


Figure 1. NIH Roadmap and research

STANDARDS required at this level: METRICS required at this level:

Public-reader

IV B. Healthcare Delivery System Needs

As much as it is an achievement to have built computer systems to support clinical care, these systems are limited in their abilities to support clinical research: From the data perspective, the data elements are too limited and the quality of the data captured often cannot be trusted for research purposes; from the user's perspective, the systems are oriented to a single patient, not to a sample or population of people across which queries need to be posed. As a result, parallel, research systems have been created. However, parallel systems lead to duplication of data entry, duplication of data monitoring, and potentials for missed events and lost data. The collection of research-oriented data must be integrated with the process of clinical care, as manifested in clinical information systems.

Clinical investigator
Basic scientist
Research administrator
Clinical manager
Public health official

Pharmaceutical companies, because they support clinical research, perforce are supporting clinical care, yet the clinical data must be re-entered and re-managed, leading to lost data or lower-quality data than they should have.

¹¹ Katz S, Patterson A, Straus S. Re-engineering the clinical research enterprise. In: NIH Roadmap for medical research: A briefing by the Director and senior staff; 2004 Feb 27, 2004; Bethesda, MD; 2004. http://nihroadmap.nih.gov/briefing/

Funders Regulators Clinician-reader Public-reader

STANDARDS required at this level: METRICS required at this level:

V A. Data, Information, and Knowledge/Repository Needs

Discussions about research infrastructure or about sharing data often devolve onto the need for common data elements or, to be more sophisticated, common vocabularies and ontologies. As is clear fromour framework (**Table 1**), such a discussion makes sense only to the degree that proposed elements, vocabularies, or ontologies support the heavy burden of function laid out in the rest of the framework.

Comparable data is the basis for the practice of evidence-based medicine. Many clinical questions will never be addressed in a randomized, blinded, controlled clinical trial. The best hope for providing clinicians answers to these questions lies with a deeper understanding of the clinical record. Recording information at the granularity with which we practice medicine, holds the promise to provide the data needed to gain an improved understanding of what constitutes the "best practice" of medicine. In order to accomplish this goal we need controlled health vocabularies to insure comparable data. Compositional vocabularies are one potential answer to the problem of providing enough content completeness to be clinically useful.

The degree to which a compositional mechanism can provide coverage for the concepts used by clinicians at the point of care is not currently known. Now using the data from the LSVT trial¹², there exists a benchmark for the coverage of these concepts by a large set of atomic and pre-coordinated concepts. By making use of standard vocabularies and adding a mechanism for creating composite terms from individual terms, we can test the value of these constructions in support of our research enterprise.

Clinical investigator
Basic scientist
Research administrator
Clinical manager
Public health official
Pharmaceutical companies
Funders
Regulators
Clinician-reader
Public-reader

STANDARDS required at this level: METRICS required at this level:

V B. Algorithmic/Tool Needs

While higher levels of our framework (

Table 1) refer to complex systems, at this level, the concern is for the component algorithms and tools needed to accomplish the higher-level functions. As in general computer sciences, algorithms are intimately tied to data structures, which is why we represent them as pairs at the same level of abstraction.

The NHII must encourage the development of automated tools to use and build compositional terminologies capable of representing a significant portion of the health concepts needed to support research and the extension of research results into the clinical practice.

Clinical investigator Basic scientist Research administrator Clinical manager

¹² Humphreys BL. McCray AT. Cheh ML. Evaluating the coverage of controlled health data terminologies: report on the results of the NLM/AHCPR large-scale vocabulary test. *Journal of the American Medical Informatics Association.* 4(6):484-500, 1997 Nov-Dec.

Public health official Pharmaceutical companies Funders Regulators Clinician-reader Public-reader

STANDARDS required at this level: METRICS required at this level:

VI. Technology

A wide range of technologies is needed to support the higher-level functions of this framework, and new technologies suggest novel ways of thinking about those, and even newer, functions. The availability of electronic data collection, whether through hand-held devices, or through desktop Web browsers has led to rethinking the clinical research process.

Clinical investigator
Basic scientist
Research administrator
Clinical manager
Public health official
Pharmaceutical companies
Funder
Regulators
Clinician-reader
Public-reader

STANDARDS required at this level: METRICS required at this level:

.

LCDR Kimberly Elenberg CCMS Program Manager Division of Human Health Sciences Food Safety Inspection Service U.S. Department of Agriculture 1400 Independence Ave., SW Aerospace Building RM 334 Washington, DC 20250-3700 (202) 690-6409

Peter L. Elkin, MD
Professor of Medicine
Director, Laboratory of Biomedical Informatics
Department of Internal Medicine
Mayo School of Medicine
Mayo Clinic, Rochester
(507) 284-1551
Fax: (507) 284-5370
http://www.mayo.edu/research/lbi

Reed M. Gardner, PhD Professor and Department Chairman Department of Medical Informatics University of Utah School of Medicine (801) 585-9428 reed.gardner@hsc.utah.edu Charles Jaffe, MD PhD Director of Medical Informatics 1800 Concord Pike, FOC CE1-316 AstraZeneca Pharmaceuticals Wilmington, DE 91850 302-885-1917

Michael G. Kahn, MD PhD Vice President, Clinical Affairs Fast Track Systems 3980 Greenbriar Blvd Boulder CO 80305 303-543-8119 mkahn@fast-track.com

Harold P. Lehmann, MD PhD Joint: Pediatrics, Health Policy and Management 2024 E Monument St, 1-201 Baltimore, MD 21287-0007 Phone 410.502.7569; 443.287.6083 Fax 410.614.2064 lehmann@jhmi.edu

Stephen Rosenfeld, MD MBA Clinical Center Chief Information Officer Associate Director for Clinical Research Information Systems NIH BLDG 10/RM 1C290 301-496-3825 stephen.rosenfeld@nih.hhs.gov